

MÓDSZERTANI TANULMÁNYOK

A KISTERÜLETI MUNKAÜGYI STATISZTIKA MÓDSZERTANA ÉS ENNEK ALKALMAZÁSA (I.)*

Kétrészes tanulmányunkban a kisterületi munkaügyi statisztikai adatok előállítására kifejlesztett, ún. kisterületi munkanélküliségi statisztikai rendszert, illetve kifejlesztése folyamatát mutatjuk be. A becslőrendszer a munkaerő-felvétel területi adatait kombinálja a regisztrált munkanélküliségi területi adatokkal megyei és megyén belüli kisterületek havi vagy negyedéves munkaügyi adatainak előállítására. A kombináció célja, hogy annak segítségével növelje az előállított adatok megbízhatóságát, csökkentve a munkaerő-felvételben a mintavételi hibából következő statisztikai hibát. Ezt a célt a becslő rendszer egyidejű adatfelvételek adatait alkalmazó, ún. kisterületi becslő függvények és az adatfelvételek múltbeli adatait is használó idősoselemző eljárások felhasználásával éri el.

TÁRGYSZÓ: Munkaügyi statisztika. Kisterületi becslőfüggvény. Idősoselemzés. Kalman-szűrő.

Az országos, a regionális és a kisterületi szintű munkanélküliségi hányadok (ráták) becslései fontos mutatói az országos, illetve a helyi gazdasági klímának. Ezért a fejlett piacgazdaságokban szakosodott szervezeteket és elemző rendszereket hoztak létre az egész gazdaságra (és részeire is) érvényes munkanélküliségi ráta (általában havi) becslésére és közzétételére. Az így előálló információt használják fel állami és önkormányzati szinten tervezési és költségvetési célokra. Ezek az információk határozzák meg a helyi foglalkoztatási és továbbképzési szolgáltatások, programok alkalmazásának szükségességét is. A kisterületi munkanélküliségi becslések segítségével döntik el az adott terület rászorultsági mértékét a központi segélyprogramokra. Természetesen e feladatok megoldásához és célok eléréséhez csak megbízható és szubjektív módon nem befolyásolható becslési módszerek és eljárások jöhetnek szóba.

A magyarországi munkaügyi információs rendszer továbbfejlesztésének is fontos része a különböző szintű területi munkanélküliségi hányadok becslésének biztosítása, beleértve a becslés elméleti és számítástechnikai kidolgozását is.

A legfontosabb igénylők és potenciális felhasználók elsősorban a kormányzati szervek, például veszélyeztetett, elmaradott területek támogatási forrásainak megállapításához vagy területfejlesztési projektek megalapozásához. De igény jelentkezik egyrészt a kisterületi

* A tanulmány szerzői: *Banai Miklós, Koleszár Kázmér, Lázár György, Lukács Béla, Prisznyák Miklós, Varga István* ezúton mondanak köszönetet *Mihályffy Lászlónak, Éltető Ödönnek* és *Christopher O'Learynek* az alapos és hasznos konzultációkért, szakmai tanácsokért és véleményekért, amelyekkel munkájukat segítették az elmúlt években. Ugyancsak köszönettel tartoznak a Központi Statisztikai Hivatalnak, ezen belül is *Lakatos Juditnak* és munkatársainak, akik szakmai konzultációk, fórumok szervezésével és az adatok gyors átadásával segítették a munkát.

munkanélküliségi rátákra a gazdasági szereplők részéről is. A korszerű beruházás-tervezés elengedhetetlen feltétele a fogadóterület állapotának foglalkoztatási szempontból való hiteles leírása, másrészt a társadalmi és érdekvédelmi szervek részéről is, hogy tevékenységüket a leghatékonyabban tudják a legfontosabb területekre koncentrálni.

Ezen igények kielégítésére kezdődött el 1993-ban, világbanki program keretében, a kisterületi munkanélküliségi statisztikai rendszer kifejlesztése. Ennek során a feladat pontos felmérése céljából, tanulmányút keretében ismertük meg az Egyesült Államok Munkaügyi Minisztériumának Statisztikai Hivatalában (Bureau of Labor Statistics – BLS) az ún. LAUS (Local Area Unemployment Statistical System) adatbecslő rendszert. Ez a rendszer havonta mintegy 6500 kisterületi munkaerő-piaci adatbecslést készít és tesz közzé. (Jellemző e kisterületi adatok iránti igény gyors növekedésére, hogy az 1993-as tanulmányutunk során a BLS még csupán 5400 kisterületi becslést készített havonta, 1994-ben viszont már 6500-ról tudósított *S. Brown*, a LAUS program vezetője.)

A cél egy olyan statisztikai elemző rendszer kidolgozása volt, amely képes a magyarországi mintegy 180 munkaügyi körzetre (a munkaügyi irodák vonzáskörzeteivel együtt), a 19 megyére és a fővárosra érvényes havi munkanélküliségi hányadok becslésének elkészítésére, a becslések statisztikai megbízhatóságának megadására, az így előálló információkból adott szempontok szerinti jelentések, összesítések „on line” rendszerű elkészítésére.

A Kisterületi Munkanélküliségi Statisztikai Rendszerben (KMSR) felhasznált módszerekről a következőket mondhatjuk.

Jó néhány ország munkaügyi információs rendszerének fontos része a kisterületi rátákat szolgáltató információs rendszer (lásd például Kanada, Egyesült Államok, Anglia, Olaszország). Az ezen országokban használt becslési eljárások jól kidolgozott, standard matematikai–statisztikai becslési módszereken alapulnak. Ezeket a módszereket adaptáljuk az általunk fejlesztett kisterületi becslési információs rendszerbe is.

Ezen módszerek sajátossága, hogy egyidejűleg becslük az aktív népesség két nagy csoportját:

- a foglalkoztatottakat,
- a munkanélkülieket.

Így a szóban forgó területekre a legaktuálisabb, legmegbízhatóbb munkanélküliségi ráta adható meg. A becslési módszer a pontosság érdekében a lehető legtöbb információt használja fel: az Országos Munkaügyi Kutató és Módszertani Központ (OMMK) regiszteradatait, továbbá a KSH lakossági és foglalkoztatásstatisztikai adatfelvételeiből származó adatokat.

Az információs rendszer arra is lehetőséget ad, hogy a foglalkoztatotti, munkanélküliségi csoportokat demográfiai, képzettségi, szociológiai szempontból is csoportosítsa és elemezze.

Befejezés előtt áll egy olyan korszerű idősorbecslő módszer kifejlesztése is, amely tovább csökkenti a becslési eredmények hibáját. Az adatok megbízhatóságának növelésére ma már ezt az ún. Kalman-szűrős módszert használják, például az Egyesült Államok munkaügyi hivatalában is.

A kisterületi munkanélküliségi statisztikai rendszer kialakítását a következő lépésekben hajtottuk végre. 1993-ban elkészítettük a „Kisterületi munkanélküliségi statisztikai rendszer kialakításának vizsgálata” című megalapozó tanulmányt. (*Kisterületi...*;1993)

A megalapozó tanulmány javaslataiból kiindulva a kisterületi munkaügyi adatbecslő és elemző rendszer kialakítását és tesztelését 1994-ben és 1995-ben három megyére és a fővárosra végzett vizsgálatunk alapján folytattuk.

1995-ben a rendszer kialakítását és tesztelését kiterjesztettük az összes megyére és az összes munkaügyi irodára. Kialakítottuk, és a megyei szinten kiértékelte kisterületi becslőfüggvények segítségével, kísérleti jellegű kiértékeléseken keresztül teszteltük a rendszer azon elemző blokkját, amely a megyei szintű becslések irodai szintre történő lebontását a megalapozó tanulmányban ismertetett leosztásos módszerek segítségével hajtja végre. A blokk kiértékeli a becslések relatív hibáját is, a megyei szintű becslések hibáira alapozva.

A rendszer kulcsfontja a megyei szintű becslések elkészítése. Erre a kisterületi becslési eljárás és a Kalman-szűrőt felhasználó idősoelemző becslési eljárás kombinált alkalmazása szolgáltatja a megoldást. A rendszer tesztváltozatába beépítettük az összes szóba jöhető becslőfüggvény-változatot, kialakítottuk a becslések varianciáját, az ún. jackknife-módszerrel meghatározó modult is. A megyei szintre adódó optimális becslőfüggvény kombináció meghatározásához elvégeztük a becslőfüggvények torzítás-vizsgálatát. Ennek eredménye alapján határoztuk meg megyénként az alkalmazható becslőfüggvények körét.

Az idősoelemző módszer alapján megyei szinten alkalmazandó eljárást specifikáltunk és azokat a rendszer tesztváltozatába beépítettük.

1996-ban már negyedévenként elkészültek a befejeződött negyedévek három hónapjára a megyénkénti és a munkaügyi irodák szerinti becslések, az 1995-ben kifejlesztett számítógépes rendszer alkalmazásával. Ezzel párhuzamosan folyt a becslési rendszer továbbfejlesztése idősoelemző módszerek felhasználásával.

1997-ben kezdtük meg a havonkénti kísérleti működtetést, az eredmények folyamatos értékelését, a számítógépes rendszer véglegesítését. Az időközben felmerült igények szerint a rendszert olyan módon volt szükséges továbbfejleszteni, hogy az képes legyen

- régiókra,
- KSH-körzetekre és
- a legalább 2000 lelket számláló településekre

is adatokat szolgáltatni a Nemzetközi Munkaügyi Szervezet (International Labour Organization – ILO) előírásainak megfelelően.

A KMSR ilyen irányú továbbfejlesztése 1998-ban indult el. A fejlesztés, a visszame-nőleges tesztelés és a kísérleti üzemeltetés 1999 végéig a következők szerint alakultak:

- módszertani továbbfejlesztés,
- a számítógépes rendszer adatbázisának kialakítása, illetve alkalmassá tétele az új feladatok ellátására,
- a településenkénti és a KSH-kistérségek szerinti ráták előállítását megvalósító programrendszer kifejlesztése és tesztelése.

A jelen tanulmány a következő fő tartalmi egységeket foglalja magába:

- a feladat megfogalmazása,
- a megalapozó tanulmány javaslatának ismertetése,
- a javaslat alapján a gyakorlati megvalósítás során létrejött statisztikai elemzőrendszer módszertani és funkcionális bemutatása az alternatív módszerek közötti választás eljárásainak ismertetésével együtt.

Az információs rendszer alkalmazásával született eredményeket egy másik tanulmányban találja az olvasó (*Banai és társai*; 2000).

A PROBLÉMA MEGFOGALMAZÁSA

Az előzőkben vázolt feladat értelmében egy adott terület munkanélküliségi rátáját kell adott időben meghatározni és ezt a feladatot időben rendszeresen elvégezni. A feladat értelmében először szükség van a munkanélküliségi ráta nemzetközi gyakorlatban alkalmazott fogalmára. A munkanélküliségi rátát (mr) a következő képlet formájában lehet felírni:

$$mr = \frac{m}{g},$$

ahol:

- m – a munkanélküliek száma,
- g – a gazdaságilag aktív népesség száma.

Adott terület munkanélküliségi rátájának adott időre vonatkozó meghatározására két lehetőség adódik:

1. közvetlenül becsüljük a hányadost az adott területre adott időben,
2. külön-külön becsüljük a hányados számlálóját és nevezőjét az adott területre az adott időben.

Mivel adott időben a gazdaságilag aktív népesség a foglalkoztatottak és a munkanélküliek összességéként áll elő, a második lehetőség további két alesetet rejt magában:

- 2a. a számláló és a nevező nagyságát közvetlenül becsüljük,
- 2b. a munkanélküliek számát (számláló) és a foglalkoztatottak számát becsüljük.

A nemzetközi gyakorlatban mindezekre a lehetőségekre dolgoztak ki módszereket és ezeket a gyakorlatban is alkalmazzák.

A feladat értelmében adott területre vonatkozóan három szintet kell megvizsgálni. Ezek:

- a) az egész ország,
- b) a főváros és a megyék,
- c) a munkaügyi körzetek (munkaügyi irodák vonzáskörzeteikkel együtt).

Az adott idő gyakorisága pedig havi, illetve éves. Az utóbbi a havi gyakorisággal előálló becslések éves szintű összegzését és összhangjuk megteremtését jelenti.

A feladat értelmében ezután

- az első lépésben megvizsgálandó, hogy az egyes lehetőségekre a nemzetközi gyakorlatban milyen becslési eljárásokat alkalmaznak, ezeknek mi az inputadat-szükséglete;
- a második lépésben megvizsgálandó, hogy Magyarországon az egyes módszerek inputadat-szükséglete rendelkezésre áll-e, ha nem, mikorra várható, hogy rendelkezésre áll;
- a harmadik lépésben megvizsgálandó, hogy az alkalmazható becslési módszerek a három megjelölt területi szint közül melyiken használhatók.

Kisterületi munkanélküliségi rátákat mind az OMMK-ban, mind a KSH-ban számolnak. Az OMMK-ban a regisztrált munkanélküliek számát tekintik a ráta számlálójaként, míg a KSH-ban a lakossági kikérdezésen alapuló munkaerő-felvétellel (MEF-adatfelvétel) az ún. ILO-kritérium szerinti munkanélküliek létszámát. Az OMMK havonta számol megyei rátákat, míg a KSH évi átlagokra egyszer. Irodai szintű rátát az OMMK-ban számolnak államigazgatási szükségletek szerint. Az OMMK-ban és a KSH-ban tehát eltérő fogalmi keretben és eltérő módszerekkel számolják a rátát.

1. Az OMMK-ban számolt kisterületi munkanélküliségi ráták definíciója a következő. A megyei ráta (mr_m):

$$mr_m = \frac{m_m}{g_{mem}}$$

ahol:

m_m – a megyei regisztrált munkanélküliség,

g_{mem} – az utolsó előtti év január 1-jei nemzetgazdasági munkaerőmérleg gazdaságilag aktív népességének leosztott része.

Az irodai szintű ráta (mr_i):

$$mr_i = \frac{m_i}{g_{nsz}}$$

ahol:

m_i – az irodai szintű regisztrált munkanélküliek száma,

g_{nsz} – az utolsó (1990-es) népszámlálás irodai szintű gazdaságilag aktív népessége.

Azonnal látszik, hogy a megyei és az irodai szintű ráták nevezője sem fogalmilag, sem időbelileg nem konzisztens. A fogalmi inkonzisztencia az adatgyűjtés természetéből és a használt fogalmi definíciók különbségéből fakad. A megyei munkanélküliségi rátában szereplő gazdaságilag aktív népesség foglalkoztatotti része nagyrészt nyilvántartási jellegű intézményi munkaügyi statisztikából megállapított létszám, kisebb részben egyéb forrásból készített becslés. (Ez utóbbi létszámba tartoznak például az intézményi statisztika megfigyelési körébe nem tartozó cégeknél foglalkoztatottak, az egyéni vállalkozók, napszámosok, alkalmi munkások, segítő családtagok stb.)

Az így megállapított létszámból arányos leosztással állítják elő a megyei szintű rátákban szereplő gazdaságilag aktív népességet.

Az 1990. január 1-jei népszámlálásból lakossági kikérdezéssel kapott adatokból számolták a gazdaságilag aktív népesség létszámát, ahol is a megkérdezettek maguk minősítették a kérdések alapján magukat foglalkoztatottnak, munkanélkülinek vagy inaktívnek.

Jól látható az is, hogy a két nevező nemcsak fogalmilag nem konzisztens, hanem a vonatkoztatási időpont is különböző. Míg a megyei ráták esetében az utolsó előtti év január 1-jén megállapított adat a nevező, addig irodai szinten az 1990-es népszámlálás változatlan adata.

Figyelembe véve, hogy a gazdaságilag aktív népesség létszáma évről évre jelentősen változik a magyar munkaerőpiacon, ez az időbeli inkonzisztencia a számított ráták megbízhatóságát erősen csökkentheti.

Vegyük először a megyei ráta számlálóját és nevezőjét. Azt kell megvizsgálnunk, hogy a használt foglalkoztatotti és munkanélküli fogalmak konzisztensek-e. Konzisztens akkor lenne a két fogalom, ha egymást kizáró lenne, egymást kiegészítő (vagyis közös gazdaságilag aktív fogalomba lehetne sorolni), kimerítő lenne (mindenki beletartozik, akinek bele kell tartoznia) és megfelelően szűk (vagyis senki nem tartozik bele, akinek nem kell). Könnyen belátható, hogy a fenti fogalompár egyik kritériumnak sem felel meg:

a) nem egymást kizáró, mivel a regisztrált munkanélküliek között megjelenhetnek olyanok, akiket például alkalmi vagy napszámos munkásként a foglalkoztatottak között figyelembe vesznek;

b) egymást nem kiegészítők, hiszen aki például elveszti munkahelyét, (és továbbra is dolgozni szeretne) nem feltétlenül jelenik meg mint regisztrált munkanélküli (nyugdíj mellett dolgozók állásvesztése); továbbá azok, akik állásvesztők, de a regisztrációból előnyük nem származik, szintén nem jelennek meg nagy valószínűséggel a regisztrált munkanélküliek között (egyéni vállalkozók, segélyezési jogosultságot nem teljesítő munkavállalók); hasonlóképpen, akik a regisztrációból kikerülnek, és aktívnak tekinthetők, nem fognak feltétlenül a foglalkoztatottak között megjelenni (a regisztrációból kikerült munkanélküliek);

c) kimerítőnek sem tekinthető, mivel nem tartalmazza azokat, akik állásvesztők vagy a munkaerőpiacra újonnan belépők, de magukat nem regisztráltatják;

d) végezetül, nem megfelelően szűk a fogalom, hiszen a gazdaságilag aktívak közé sorol olyanokat is, akik valójában nem szereplői a munkaerőpiacnak; azokról a személyekről van szó, akiket regisztráltak ugyan, de nem kívánnak dolgozni, munkát nem keresnek.

2. A KSH MEF-adatfelvétele a ráta számlálóját és nevezőjét tekintve fogalmilag és a vonatkoztatási időpont szempontjából is konzisztens. Kielégíti a kölcsönös kizárás, a kiegészítés, a lefedettség és a megfelelő szűkösség kritériumát. Továbbá, mivel azonos időpontra vonatkozó létszám a számláló és a nevező értéke is, így az időbeli inkonzisztencia is fennáll. Probléma viszont, hogy az adatfelvétel a minta nagysága miatt nem alkalmas arra, hogy módszertanilag megalapozott, statisztikailag megbízható becslést adjon a megyei és az irodai szintű munkanélküliségi rátákra havonta, sőt jelenleg még negyedévente sem. A második megoldandó feladat tehát a fogalmilag és idő szerint konzisztens mérések megbízhatóságának növelése.

Az eddigiek egyértelműen mutatják a feladatot: fogalmilag konzisztens havi, esetleg negyedéves szinten statisztikailag megbízható munkanélküliségi rátákat kell becsülni.

JAVASLAT A KISTERÜLETI MUNKANÉLKÜLISÉGI STATISZTIKAI RENDSZER KIALAKÍTÁSÁRA

A probléma megfogalmazásából egyértelmű, hogy a javasolt megoldásnak a következőket kell tisztáznia.

A feladat

1. A havi adatbecslések előállításához, országos szinten a jelenlegi negyedéves becslések helyett havi becslésekre van szükség. Ezt a területi szintű idősoros kiigazítás (benchmark) igényli.

2. A megyei szintű, adott megbízhatósági szintet teljesítő havi becslések előállítása. Módszertanilag ez a feladat kulcsfontosságú része.

3. Havi gyakoriságú irodai szintű becslések előállítása az adott területre vonatkozó adatokból közvetlenül, vagy a megyei szintű és a helyi adatok kombinálásával.

4. A megoldásnak a rendelkezésre álló adatokra kell támaszkodnia.

Adatforrások

A rendelkezésre álló adatok közül a KSH-ból származók:

- a munkaerő-felvétel (MEF) anonimizált személyi szintű és területi szintű aggregált havi adatai 1992. januártól,
- az 1990-es népszámlálás település szintű adatai és ezek évenkénti továbbvezetései,
- a munkaügyi statisztikai adatok havonta, illetve negyedévente 1992. januártól.

A KSH megyei igazgatóságaitól beszerezhetők:

- a területi munkaügyi adatok negyedévente, 1992. I. negyedévtől.

Az OMMK-ból származók:

- az anonimizált személyi szintű regisztrált munkanélküliségi adatok havonta, 1990. januártól.

Módszerek

Ezeket az adatokat figyelembe véve és a nemzetközi gyakorlatra támaszkodva a módszerek következő két nagy osztálya került be a javaslatunkba:

- a) az adott területre az adott időpontra vonatkozó, azaz egyidejű adatokat használó becslési eljárások,
- b) az adott területre az adott időpontban és az azt megelőző időpontokban rendelkezésre álló adatokat, azaz idősorokat használó becslési eljárások.

Az egyidejű adatokat alkalmazó becslési eljárásokból megvizsgáltuk

- az ún. Handbook-módszer adaptációját és
- az ún. kisterületi becslési módszereket.

Az előbbi a BLS-ben alkalmazzák a munkaerő-piaci körzetekre (labor market areas) vonatkozó becslések előállítására. A gyakorlatban a módszer alkalmazására mind ez idáig még nem került sor.

Az utóbbi módszereket akkor alkalmazzák, ha az adott területre mintavételen alapuló közvetlen becsléssel nem származtathatók elfogadható megbízhatóságú adatok. A módszer szerint a mintavételes becslés megbízhatóságának javítására külső adatokat kell bevonni a becslésbe. Ezeket a becslési eljárásokat alkalmazzák például Kanadában kisterületi becslések származtatására.

A módszerek diszkussziója

Tesztelésre a becslési eljárások két csoportját javasoltuk: a hányados-becsléseket és a korrekciós becsléseket.

Az előbbieket a becslés mennyiségét (munkanélküliek vagy foglalkoztatottak) az MEF-ből az adott területre (egyszerű becslés) vagy az adott területet magába foglaló nagyobb területre (szintetikus becslés) kapott becslési értékek arányában állítják elő a bevont külső mennyiségből (például munkanélküliség esetén ez a mennyiség az adott terület regisztrált munkanélkülieinek a száma). De hányados-becslés az Egyesült Államokban alkalmazott ún. népességrészesedési és lakosságiigény-módszerek is (az utóbbi elnevezés a munkanélküliségi ellátás igénylésére utal). Ezek a nagyobb területre vonatkozó becslési értéket a kisebb területek lakónépességei vagy az azokon regisztrált munkanélkülieinek arányában osztják szét.

A korrekciós becslések az adott területre vonatkozó MEF-becslést korrekciós tag hozzáadásával javítják. A korrekciós tagok a MEF-ből származó mennyiségeket (ezek lehetnek egyszerű vagy szintetikus típusúak) és külső adatokat (regisztrált munkanélküliek száma az adott területen) tartalmaznak.

A becslő rendszer tesztváltozatába, a rétegzett változatokkal együtt, mintegy 26-féle becslőfüggvényt és a lakosságiigény-, népességrészesedési módszereket építettünk be és teszteltünk. A tesztelés során nem csupán az elérhető varianciajavulást vizsgáltuk, hanem az egyes becslési eljárásokban fellelhető torzító hatásokat is figyeltük. E célból a becslő rendszerbe a torzításvizsgálatok elvégzését lehetővé tevő eljárásokat alkalmaztunk. A becslőfüggvények varianciájának kiértékelésére a jackknife-eljárást használtuk.

A MEF-becslés varianciájának csökkentésére alkalmazott kisterületi becslési eljárások elvi szempontból az eredményt különböző irányokba korrigálhatják. Teljes körű felmérés hiányában nem lehet megállapítani, hogy melyik eljárás adja az igazi értékhez legközelebb álló becslést. A megoldást az összes lehetséges becslőfüggvény súlyozott átlaga szolgáltatná, ahol a méréselméleteknek megfelelően, a súlyokat a varianciák inverzei adják.

Jelen esetben azonban ez az eljárás súlyos hibákhoz vezethet, hiszen a becslőfüggvények függenek egymástól, mivel csak a korrekciós tagok vagy faktorok különböznek. Ezért megvan annak a veszélye, hogy kiegyensúlyozatlan számú becslőfüggvény-kombinációval a súlyozott átlag is kiegyensúlyozatlan (instabil) lesz. Ennek a problémának a megoldására a súlyozott átlag olyan általánosítására van szükség, amely automatikusan kezeli a becslőfüggvények erős korrelációit.

Ezt az eredményt a becslőfüggvények azon centruma szolgáltatja, amelyet az ingadozások általánosított minimum elvének megfelelően képezünk. Az eljárást részletesen egy korábbi tanulmányunkban (*Kisterületi...*; 1996a) mutattuk be. Ennek lényege a következő.

Jelölje az i -edik becslőfüggvényt h_i , a centrumot pedig h . A becslőfüggvények kovarianciamátrixa ekkor:

$$\sigma_{ik} = E(h_i h_k) - E(h_i)E(h_k).$$

Ezután annak megkövetelésével, hogy a fluktuációk a centrum körül minimálisak legyenek, kapjuk:

$$\sum_{ik} \sigma^{ik} (h_i - h)(h_k - h) = \text{minimum},$$

ahol σ^{ik} jelöli az inverz kovarianciamátrix megfelelő elemét. Ebből az összefüggésből megkaphatjuk a h centrumot és annak varianciáját (*Kisterületi...*; 1996a).

Látható, hogy két erősen korreláló becslőfüggvény esetére σ^{ik} értéke garantálja, hogy azok teljes centrumbeli súlya csupán kismértékben legyen magasabb, mint bármelyiké külön-külön. Ezért az eljárás korrigálja a hasonló becslőfüggvények egyidejű alkalmazása során előforduló hibát. A becslőrendszerbe beépítettük a centrumnak és varianciájának kiértékelésére szolgáló eljárásokat is.

Az idősorokat alkalmazó becslési eljárások közül a jel+zaj típusú idősorelemző eljárást építettük be a tesztrendszerbe, ahol a jel képviseli a keresett értéket (populációs érték), a zaj pedig a mintavételes eljárás hibáját írja le. A zaj elválasztását ebben a módszerben az ún. Kalman-szűrő algoritmus végzi. A jelet magyarázó változót(kat) alkalmazó regressziós tagból, trendből és szezonális ingadozásokat megőrző tagból felépített strukturált idősormodell, a zajt pedig irreguláris változásokat, kiugró értékeket és a megfigyelt idősorhoz tartozó hibaidősört leíró, ún. ARMA-modell együttese írja le. A módszer alkalmazhatósága függ attól, hogy a rendelkezésre álló bemeneti idősorok elég hosszúak-e.

Ebben a becslési eljárásban a jel (*Jel*) adott t időpontra vonatkozó becslése a korábbi időpontokra vonatkozó jelértékek egy lépéses előrejelzéséből (*Jél*), a javítandó t időpontra vonatkozó becslési értékből (*Becs*) és a mintavételi hibát leíró zaj modellalakjából (*Z*) a t időpontra kapott értékből a következő összefüggéssel állítható elő:

$$Jel(t) = (1-w(t))Jél(t) + w(t)(Becs(t) - Z(t)),$$

ahol w a Kalman-szűrő által előállított, 0 és 1 közötti értékeket felvevő súly. Értéke a jel és a zaj varianciáinak hányadosától függ oly módon, hogy minél nagyobb a zaj varianciája annál kisebb ez az érték és fordítva. Tehát kis mintavételi hiba esetén az aktuális becslés, míg nagy mintavételi hiba esetén az előrejelzett jel dominál az aktuális jel értékében (az eljárás teljesíti a mintavételi tervvel való konzisztenciát). Ilyen módon látható az összefüggésből, hogyan szűri ki az eljárás az eredeti becslés hibáját (a mintavételi hibát) a korrigált becslésből.

Területi szinten a becslések előállításánál biztosítani kell az additivitás elvének teljesülését, azaz a kisebb területekre kapott becslések összegének ki kell adnia a kisterületet magában foglaló nagyobb területre vonatkozó becslés értékét. Emellett a havi becslések és az éves becslések összhangját is biztosítani kell.

A javasolt információs rendszer

Javaslatunk a területek hierarchiájában a következőképpen fogalmazódott meg.

1. *Havi becslések az ország egészére.* Ezt a becslést mind a munkanélküliségre, mind a foglalkoztatottságra a MEF-havi becsléseire kifejlesztett idősor-regressziós modellekkel állíthatjuk elő. Ezen modellek bemeneti adatai egyrészt a MEF havi munkanélküliségi és foglalkoztatási adatbecsléseinek idősorai, másrészt a havi és negyedéves foglalkoztatásstatisztikai adatsorok, a népesség-továbbvezetés adatsorai és a regisztrált munkanélküliek havi adatsorai magyarázó változók előállításához. A munkanélküliségi hányadot magából a munkanélküliség becslésére kifejlesztett modelltől kapjuk, mivel azt célszerű közvetlenül a rátára kialakítani és a létszámadat pedig a rátából és a foglalkoztatottak számának becsléséből levezetni. Ekkor elméleti megfontolások alapján ezen becslések

statisztikai hibáinak kisebbnek kell lenniük a MEF-becslések hibáinál. A havi becsléseknek összhangba kell lenniük a KSH által kapott éves becslési eredményekkel. Ezt a Denton-féle kiigazítási (benchmark) eljárás alkalmazásával biztosítjuk.

2. *Havi becslések a fővárosra és a megyékre.* Ezen a területi szinten javaslatunk a kisterületi becselőfüggvényekre alapozott centrum és az idősolelemző eljárás kombinációjából áll. Első lépésben az adott területre adódó centrum segítségével meghatározzuk a becsléseket mind a foglalkoztatottságra és a munkanélküliségre, mind a munkanélküliségi hányadra, 1992 januárjáig visszamenőleg. Második lépésben az első lépésből kapott idősorokra mint bemeneti idősorokra alkalmazzuk az idősor-regressziós eljárást. Ily módon a centrum alkalmazásával elért varianciacsökkenést tovább javítjuk a Kalman-szűrőt alkalmazó jelkiszűrő eljárással. A bemeneti adatok ebben az esetben is az egyes területekre vonatkozó MEF havi munkanélküliségi és foglalkoztatási adatbecslései, illetve ezek idősorai, a havi és negyedéves foglalkoztatásstatisztikai adatok, illetve adatsorok (a székhely szerint gyűjtött adatokat valamilyen módszerrel szét kell osztani a megyékre és a fővárosra); a népesség-továbbvezetés adatai, illetve adatsorai az egyes területekre és a regisztrált munkanélküliek havi adatai, illetve adatsorai az egyes területekre, mind a korrekciós tagok változóihoz, mind a magyarázó változók előállításához.

A idősoros becslés során a munkanélküliségi hányadot magából a munkanélküliség becslésére kifejlesztett modelltől kapjuk, miután azt célszerű ebben az esetben is közvetlenül a rátára kialakítani, a létszámadatot pedig a rátából és a foglalkoztatottak számának becsléséből levezetni.

Az additivitás elvének teljesüléséhez biztosítani kell, hogy a területekre kapott becslések összegei kiadják az országos szinten kapott havi becslési értékeket. A havi becsléseknek az egyes területeken is összhangban kell lenniük a KSH által az adott területre kapott éves becslési eredményekkel. Ezt – akárcsak országos szinten – most is a Denton-féle eljárás alkalmazásával biztosítjuk.

3. *Havi becslések az egyes munkaügyi irodák területére.* Ezen a területi szinten javaslatunk szintén kombinált. Egyik alternatíva: minden irodára alkalmazható a lakosságiigény-módszer adaptált változata, amely a megyei adatok leosztását jelenti. *Foglalkoztatási adatbecslés* esetén a megyei adatot az egyes irodák területére vonatkozó népszámláláskori foglalkoztatotti létszámadatok és az aktuális hónapra vonatkozó továbbvezetett népességadatok felhasználásával osztjuk szét. *Munkanélküliségi adatbecslés* esetén a megyei adatot az egyes irodák területére vonatkozó (korcsoportos bontású) regisztrált munkanélküliségi és továbbvezetett népességadatok felhasználásával osztjuk szét. Ebben az esetben a becslések hibáját a megyei adatok hibájából a hibaterjedés módszerével határozzuk meg.

A másik alternatívát a kisterületi becselőfüggvények adják. Minden iroda területére, ahol az adott megyében alkalmazható becselőfüggvények közül találunk olyanokat, amelyek időben stabilan kiértékelhetők az adott területre, azokat kiértékeljük. Statisztikai hibájukat a segítségükkel előállított becslési értékek idősolelemzésével határozzuk meg (a jackknife-módszer, néhány irodát leszámítva, irodai szinten nem alkalmazható) oly módon, hogy az X-11 idősolelemző eljárás segítségével az idősort trend-ciklus-, szezonális és irreguláris összetevőkre bontjuk és a statisztikai ingadozást a szisztematikus összetevők eltávolításával határozzuk meg. Ezután az adott irodára azt a becslést javasoljuk publikálni, amelynek a megbízhatósági mutatója jobbnak adódik. Természetesen a

munkanélküliségi és a foglalkoztatottsági adatokat egy adott időpontban ugyanazzal a módszerrel (leosztással vagy kisterületi becsléssel) kell előállítani. A munkanélküliségi rátát pedig a létszámadatokból a

$$\text{MUNKANÉLKÜLIEK}/(\text{MUNKANÉLKÜLIEK} + \text{FOGLALKOZTATOTTAK})$$

összefüggés segítségével állítjuk elő az adott hónapra, amelynek hibája azután vagy hibaterjedéssel vagy idősoelemzéssel határozható meg.

A területi szintű additivitást a megyei adatoknak az országos adatokhoz való igazításával lehet elérni, a havi adatok éves adatokkal való összhangját pedig a korábban már említett Denton-féle eljárással biztosítjuk, ahol az éves adatokat a megyei szintű éves adatok leosztásával állítjuk elő (idősoelemzésre nem áll rendelkezésre elegendő visszamenőleges adat).

(A tanulmány II., befejező részét a *Statistikai Szemle* következő száma közli.)